

What is claimed is:

- 1 1. A computer-implemented method of analyzing linguistic terms, the method
2 comprising:
3 (a) scanning a plurality of documents for variants of a linguistic term;
4 and
5 (b) tracking relative occurrences of a plurality of variants of the
6 linguistic term found in the plurality of documents during scanning to
7 determine an acceptable usage of the linguistic term.
- 1 2. The method of claim 1, further comprising retrieving the plurality of
2 documents from a network, wherein scanning the plurality of documents includes
3 scanning each document subsequent to retrieval of the document from the network.
- 1 3. The method of claim 2, wherein retrieving the plurality of documents from
2 the network comprises retrieving the plurality of documents from at least one Internet
3 web site in response to a user browsing the at least one Internet web site, and wherein
4 scanning the plurality of documents includes scanning each document upon retrieval
5 of that document from the at least one Internet web site.
- 1 4. The method of claim 2, further comprising determining whether a retrieved
2 document has already been scanned before scanning the retrieved document.
- 1 5. The method of claim 2, further comprising determining whether to scan a
2 retrieved document based upon a source parameter associated with the linguistic term.
- 1 6. The method of claim 1, further comprising browsing a second plurality of
2 documents retrieved from at least one Internet web site in response to user input,
3 wherein scanning the first plurality of documents is performed concurrently with
4 browsing the second plurality of documents.

1 7. The method of claim 6, wherein scanning the first plurality of documents is
2 performed in a background thread while documents from the second plurality of
3 documents are being browsed.

1 8. The method of claim 7, wherein scanning the first plurality of documents
2 includes scanning documents stored in a local history cache.

1 9. The method of claim 1, wherein the linguistic term comprises a single
2 word.

1 10. The method of claim 1, wherein the linguistic term comprises a phrase.

1 11. The method of claim 1, wherein the linguistic term comprises an acronym.

1 12. The method of claim 1, wherein the plurality of variants differ from one
2 another based upon at least one of punctuation, spelling, capitalization, hyphenation,
3 and definition.

1 13. The method of claim 1, wherein scanning the plurality of documents
2 includes scanning a document for an enumerated variant of the linguistic term.

1 14. The method of claim 1, wherein scanning the plurality of documents
2 includes scanning a document for an unenumerated variant of the linguistic term.

1 15. The method of claim 14, wherein scanning the document for the
2 unenumerated variant of the linguistic term includes scanning the document using
3 phonetic comparison.

1 16. The method of claim 1, wherein tracking relative occurrences of the
2 plurality of variants includes weighting occurrences based upon locations of such
3 occurrences within the plurality of documents.

1 17. The method of claim 1, wherein tracking relative occurrences of the
2 plurality of variants includes weighting occurrences based upon document types of the
3 documents within which such occurrences are found.

1 18. The method of claim 1, further comprising storing a variant of the
2 linguistic term in an electronic dictionary.

1 19. The method of claim 18, further comprising spell checking a document
2 using the electronic dictionary subsequent to storing the variant in the electronic
3 dictionary.

1 20. The method of claim 1, wherein tracking relative occurrences of the
2 plurality of variants includes storing context information associated with each
3 occurrence of a variant of the linguistic term.

1 21. The method of claim 1, wherein scanning the plurality of documents
2 includes scanning a document for a spell definition tag that identifies a variant of the
3 linguistic term.

1 22. The method of claim 1, wherein scanning the plurality of documents and
2 tracking relative occurrences are performed responsive to detecting a variant of the
3 linguistic term during spell checking of a document.

1. The first part of the document is a list of references. The references are listed in a standard format, with the author's name, the title of the work, and the publisher. The references are as follows:

1. The first part of the document is a list of references. The references are listed in a standard format, with the author's name, the title of the work, and the publisher. The references are as follows:

1 24. An apparatus, comprising:
2 (a) a memory; and
3 (b) a program resident in the memory and configured to determine an
4 acceptable usage of a linguistic term by scanning a plurality of documents for
5 variants of the linguistic term and tracking relative occurrences of a plurality
6 of variants of the linguistic term found in the plurality of documents during
7 scanning.

1 25. The apparatus of claim 24, wherein the program is further configured to
2 retrieve the plurality of documents from at least one Internet web site in response to a
3 user browsing the at least one Internet web site and scan the plurality of documents by
4 scanning each document upon retrieval of that document from the at least one Internet
5 web site.

1 26. The apparatus of claim 25, wherein the program is further configured to
2 determine whether a retrieved document has already been scanned before scanning the
3 retrieved document.

1 27. The apparatus of claim 25, wherein the program is further configured to
2 determine whether to scan a retrieved document based upon a source parameter
3 associated with the linguistic term.

1 28. The apparatus of claim 24, wherein the program is further configured to
2 browse a second plurality of documents retrieved from at least one Internet web site in
3 response to user input, and scan the first plurality of documents concurrently with
4 browsing the second plurality of documents.

1 29. The apparatus of claim 24, wherein the linguistic term is selected from the
2 group consisting of a single word, a phrase, and an acronym.

1 30. The apparatus of claim 24, wherein the plurality of variants differ from
2 one another based upon at least one of punctuation, spelling, capitalization,
3 hyphenation, and definition.

1 31. The apparatus of claim 24, wherein the program is configured to scan the
2 plurality of documents by scanning a document for an enumerated variant of the
3 linguistic term.

1 32. The apparatus of claim 24, wherein the program is configured to scan the
2 plurality of documents by scanning a document for an unenumerated variant of the
3 linguistic term.

1 33. The apparatus of claim 24, wherein the program is configured to track
2 relative occurrences of the plurality of variants by weighting occurrences based upon
3 at least one of locations of such occurrences within the plurality of documents, and
4 document types of the documents within which such occurrences are found.

1 34. The apparatus of claim 26, wherein the program is further configured to
2 store a variant of the linguistic term in an electronic dictionary, the apparatus further
3 comprising a spell checker configured to spell check a document using the electronic
4 dictionary subsequent to the variant being stored in the electronic dictionary.

1 35. The apparatus of claim 26, wherein the program is further configured to
2 store context information associated with each occurrence of a variant of the linguistic
3 term.

1 36. The apparatus of claim 26, wherein the program is configured to scan a
2 document for a spell definition tag that identifies a variant of the linguistic term.

- 1 37. A program product, comprising:
- 2 (a) a program configured to determine an acceptable usage of a
- 3 linguistic term by scanning a plurality of documents for variants of the
- 4 linguistic term and tracking relative occurrences of a plurality of variants of
- 5 the linguistic term found in the plurality of documents during scanning; and
- 6 (b) a signal bearing medium bearing the program.

- 1 38. The program product of claim 37, wherein the signal bearing medium
- 2 includes at least one of a transmission medium and a recordable medium.

IBM ROC9-2000-0191
WH&E IBM/166

1 39. A program product, comprising:

2 (a) a document, the document including a tag that identifies an
3 acceptable variant of a linguistic term and a definition of the linguistic term;
4 and

5 (b) a signal bearing medium bearing the document.

IBM CORPORATION
ARMONK, NEW YORK 10504
U.S. PATENT AND
TRADEMARK OFFICE
WASHINGTON, D.C. 20540

1 40. A method of spell checking a document, the method comprising:
2 (a) comparing terms in a first document against an electronic
3 dictionary; and
4 (b) in response to determining during the comparison that a term from
5 the document is not in the electronic dictionary, automatically scanning a
6 plurality of documents from the Internet to identify at least one acceptable
7 usage of the term.

1 41. The method of claim 40, further comprising:
2 (a) tracking relative occurrences of a plurality of variants of the term
3 found in the plurality of documents; and
4 (b) displaying results of such tracking to a user.

1 42. A method of managing an electronic dictionary, the method comprising:
2 (a) detecting a spell definition tag within a document retrieved from
3 the Internet that identifies an acceptable variant of a linguistic term; and
4 (b) in response to detecting the spell definition tag, automatically
5 adding the acceptable variant of the linguistic term to an electronic dictionary.

1 43. The method of claim 42, wherein detecting the spell definition tag is
2 performed during user browsing of the Internet.

IBM ROC9-2000-0191
WH&E IBM/166